OPTIMAL JOINT MEASUREMENTS OF POSITION AND MOMENTUM

D M APPLEBY
Department of Physics, Queen Mary and Westfield College, Mile End Rd, London E1 4NS, UK

(e-mail: D.M.Appleby@qmw.ac.uk)

Abstract

The distribution of measured values for maximally accurate, unbiased simultaneous measurements of position and momentum is investigated. It is shown, that if the measurement is retrodictively optimal, then the distribution of results is given by the initial state Husimi function (or Q-representation). If the measurement is predictively optimal, then the distribution of results is related to the final state anti-Husimi function (or P-representation). The significance of this universal property for the interpretation of the Husimi function is discussed.

PACS number: 03.65.Bz

Report no. QMW-PH-98-14

1. Introduction

There is currently some interest in simultaneous measurements of position and momentum [1, 2, 3, 4, 5, 6, 7, 8, 9]. Measurements of this kind have an immediate, technical relevance to the field of quantum optics. They also have a rather more general, conceptual relevance to the problem of understanding the classical limit.

In two previous papers [10, 11] we discussed the accuracy of such measurements. We began with Braginsky and Khalili's analysis [12] of single measurements of x only, and extended it to the case of simultaneous measurements of x and p together. We identified two types of error: the retrodictive (or determinative) errors $\Delta_{\mathrm{ei}}x$, $\Delta_{\mathrm{ei}}p$; and the predictive (or preparative) errors $\Delta_{\mathrm{ef}}x$, $\Delta_{\mathrm{ef}}p$. We showed, that subject to some rather unrestrictive assumptions regarding the nature of the measurement process, they satisfy the retrodictive error relationship

$$\Delta_{\mathrm{ei}}x \, \Delta_{\mathrm{ei}}p \geq \frac{\hbar}{2}$$

and the predictive error relationship

$$\Delta_{\mathrm{ef}}x \, \Delta_{\mathrm{ef}}p \geq \frac{\hbar}{2}$$

In the following we address the question: what (if anything) can be said about the distribution of measured values in those cases where the lower bound set by one of these inequalities is actually achieved?

We begin, in Section 2, by considering measurements which are retrodictively optimal. We define a retrodictively optimal measurement to be any measurement belonging to the class of processes defined in ref. [11] which minimises the product of retrodictive errors (so that $\Delta_{\mathrm{ei}}x \, \Delta_{\mathrm{ei}}p = \frac{\hbar}{2}$), and which is retrodictively unbiased [so that the systematic errors of retrodiction are zero—see Eq. (1) below]. We show, that for such measurements, the distribution of measured values is always given by the initial system state Husimi function [13, 14]. This result is the extension, to the general class of measurement processes defined in ref. [11], of the result proved by Ali and Prugovečki [9, 15] for the case of measurement processes which are Galilean covariant, and (using rather different methods) in ref. [16] for the particular case of the Arthurs-Kelly process.

A number of related results have been obtained by other authors. In the case of the Arthurs-Kelly process several authors [1, 8] have shown, that the Husimi function describes the distribution of measured values for certain choices of initial apparatus state. Leonhardt and Paul [3] have shown that the same is true for a number of other processes. However, these authors all confine themselves to particular examples of simultaneous measurement processes. They do not consider measurement processes in general. Moreover, they do not relate the distribution of measured values to the accuracy of the measurement process. In particular, they do not show that the Husimi function describes the distribution of results in precisely those cases where the measurement is retrodictively "optimal" or "best".

Wódkiewicz has proposed an operational approach to the problem of phase space measurement [6, 7]. If one takes the filter reference state (or "quantum ruler") used to define his operational distribution to be a squeezed vacuum state, and a minimum uncertainty state for $\hat{x}$ and $\hat{p}$, then one obtains the Husimi function. It could be said that the Husimi function is the operational distribution corresponding to the

case when the quantum ruler is most exactly and finely calibrated—a fact which obviously ties in with the result which we prove in Section 2 below.

However, the result which is most similar to ours is the one obtained by Ali and Prugovečki [9, 15], working within the framework of the approach based on POVM's (positive operator valued measures) and unsharp observables. In fact, their result is the same as ours, except that we prove it under much less restrictive conditions (unlike Ali and Prugovečki we do not assume Galilean covariance. Galilean covariance is a consequence of the result which we prove, not a presupposition). It may also be worth remarking that our way of analysing the concept of a simultaneous measurement process is rather different from theirs. In particular, the objections recently raised by Uffink [17] do not apply to our arguments.

In Section 3 we go on to consider predictively optimal measurements—*i.e.* measurements of the type defined in ref. [11] which minimise the product of predictive errors (so that $\Delta_{\rm ef} x \, \Delta_{\rm ef} p = \frac{\hbar}{2}$). We show, that in the case of such a measurement, the distribution of results is related to the final state anti-Husimi function [14, 18] (the P-function of quantum optics). This result also represents an extension, to the general class of measurement processes defined in ref. [11], of a result proved in ref. [16], for the special case of the Arthurs-Kelly process.

In Section 4 we conclude by discussing the bearing of our results on the interpretation of the Husimi function. In Section 2 we show that the Husimi function describes the outcome of *any* retrodictively optimal process. In other words, the Husimi function has a universal significance. We will argue that this lends some support to the idea, that the Husimi function is the quantum mechanical entity which most nearly resembles the classical concept, of the "real" or "objective" distribution describing an ensemble of identically prepared systems.

2. Retrodictively Optimal Measurements

We will say that a simultaneous measurement process of the kind defined in ref. [11] is *retrodictively optimal* if

1. The process is retrodictively unbiased, so that
$$\langle \psi \otimes \phi_{\rm ap} | \, \hat{\epsilon}_{\rm Xi} \, | \psi \otimes \phi_{\rm ap} \rangle = \langle \psi \otimes \phi_{\rm ap} | \, \hat{\epsilon}_{\rm Pi} \, | \psi \otimes \phi_{\rm ap} \rangle = 0 \tag{1}$$
for all $|\psi\rangle \in \mathcal{H}_{\rm sy}$.
2. The product of retrodictive errors achieves its lower bound, so that
$$\Delta_{\rm ei} x \, \Delta_{\rm ei} p = \frac{\hbar}{2} \tag{2}$$

Here and in the sequel we employ the notation and terminology of ref. [11]. Thus, $|\psi\rangle \in \mathcal{H}_{\rm sy}$ and $|\phi_{\rm ap}\rangle \in \mathcal{H}_{\rm ap}$ are the initial states of the system and apparatus respectively. $\hat{\epsilon}_{\rm Xi}$, $\hat{\epsilon}_{\rm Pi}$ are the retrodictive error operators. $\Delta_{\rm ei} x$, $\Delta_{\rm ei} p$ are the maximal rms errors of retrodiction.

In ref [16] we considered the special case of the Arthurs-Kelly process. In that case one has the commutation relation
$$[\hat{\epsilon}_{\rm Xi}, \hat{\epsilon}_{\rm Pi}] = -i\hbar \tag{3}$$
This relationship, and the condition of Eq. (2), together imply Eq. (1). In the general case, however, it is necessary to impose the requirement, that the measurement be retrodictively unbiased, as a separate condition.

In the general case the case the commutation relationship of Eq. (3) cannot be assumed. However, it was shown in ref. [11] that Eq. (1) implies the weaker statement

$$\langle \psi \otimes \phi_{\mathrm{ap}} | \, [\hat{\epsilon}_{\mathrm{Xi}}, \hat{\epsilon}_{\mathrm{Pi}}] \, | \psi \otimes \phi_{\mathrm{ap}} \rangle = -i\hbar \tag{4}$$

for every normalised $|\psi\rangle \in \mathcal{H}_{\mathrm{sy}}$ [but fixed $|\phi_{\mathrm{ap}}\rangle$]. It turns out that this is enough to prove, that the distribution of measured values is given by the initial system state Husimi function, for any retrodictively optimal process. However, the fact that we can no longer assume the commutation relationship of Eq. (3), means that the proof of this statement is less straightforward than the proof given in ref. [16], for the special case of the Arthurs-Kelly process.

In view of Eqs. (2) and (4) we have

$$\left\langle \psi \otimes \phi_{\mathrm{ap}} \right| \hat{\epsilon}^2_{\mathrm{Xi}} \left| \psi \otimes \phi_{\mathrm{ap}} \right\rangle \left\langle \psi \otimes \phi_{\mathrm{ap}} \right| \hat{\epsilon}^2_{\mathrm{Pi}} \left| \psi \otimes \phi_{\mathrm{ap}} \right\rangle = \frac{\hbar^2}{4} \tag{5}$$

for every normalised $|\psi\rangle \in \mathcal{H}_{\mathrm{sy}}$. We deduce:

Lemma 1. *Given any retrodictively optimal measurement process with initial apparatus state $|\phi_{\mathrm{ap}}\rangle$, there exists a fixed number λ_{i} such that*

$$\left\langle \psi \otimes \phi_{\mathrm{ap}} \right| \hat{\epsilon}^2_{\mathrm{Xi}} \left| \psi \otimes \phi_{\mathrm{ap}} \right\rangle = \frac{\lambda_{\mathrm{i}}^2}{2}$$
$$\left\langle \psi \otimes \phi_{\mathrm{ap}} \right| \hat{\epsilon}^2_{\mathrm{Pi}} \left| \psi \otimes \phi_{\mathrm{ap}} \right\rangle = \frac{\hbar^2}{2\lambda_{\mathrm{i}}^2}$$

for every normalised $|\psi\rangle \in \mathcal{H}_{\mathrm{sy}}$.

Remark. We will refer to λ_{i} as the retrodictive spatial resolution of the measurement.

Proof. For each normalised $|\psi\rangle \in \mathcal{H}_{\mathrm{sy}}$ define the number λ_ψ by

$$\lambda_\psi = \left(2 \left\langle \psi \otimes \phi_{\mathrm{ap}} \right| \hat{\epsilon}^2_{\mathrm{Xi}} \left| \psi \otimes \phi_{\mathrm{ap}} \right\rangle\right)^{\frac{1}{2}}$$

In view of Eq. (5) we then have

$$\left(\left\langle \psi \otimes \phi_{\mathrm{ap}} \right| \hat{\epsilon}^2_{\mathrm{Pi}} \left| \psi \otimes \phi_{\mathrm{ap}} \right\rangle\right)^{\frac{1}{2}} = \frac{\hbar}{\sqrt{2}\lambda_\psi}$$

We have from the definitions [11] of $\Delta_{\mathrm{ei}} x$, $\Delta_{\mathrm{ei}} p$

$$\Delta_{\mathrm{ei}} x = \sup_{|\psi\rangle \in \mathcal{S}} \left(\left\langle \psi \otimes \phi_{\mathrm{ap}} \right| \hat{\epsilon}^2_{\mathrm{Xi}} \left| \psi \otimes \phi_{\mathrm{ap}} \right\rangle\right)^{\frac{1}{2}} = \frac{\sup_{|\psi\rangle \in \mathcal{S}} (\lambda_\psi)}{\sqrt{2}}$$

and

$$\Delta_{\mathrm{ei}} p = \sup_{|\psi\rangle \in \mathcal{S}} \left(\left\langle \psi \otimes \phi_{\mathrm{ap}} \right| \hat{\epsilon}^2_{\mathrm{Pi}} \left| \psi \otimes \phi_{\mathrm{ap}} \right\rangle\right)^{\frac{1}{2}} = \frac{\hbar}{\sqrt{2} \inf_{|\psi\rangle \in \mathcal{S}} (\lambda_\psi)}$$

where $\mathcal{S}$ denotes the unit sphere in the system state space. In view of Eq. (2) it then follows

$$\inf_{|\psi\rangle \in \mathcal{S}} (\lambda_\psi) = \sup_{|\psi\rangle \in \mathcal{S}} (\lambda_\psi)$$

which means that λ_ψ must be constant. $\square$

We next define the operators

$$\hat{c}_{\lambda_\mathrm{i}} = \frac{1}{\sqrt{2}}\left(\frac{1}{\lambda_\mathrm{i}}\hat{\epsilon}_\mathrm{Xi} - \frac{i\lambda_\mathrm{i}}{\hbar}\hat{\epsilon}_\mathrm{Pi}\right)$$
$$\hat{c}^\dagger_{\lambda_\mathrm{i}} = \frac{1}{\sqrt{2}}\left(\frac{1}{\lambda_\mathrm{i}}\hat{\epsilon}_\mathrm{Xi} + \frac{i\lambda_\mathrm{i}}{\hbar}\hat{\epsilon}_\mathrm{Pi}\right) \tag{6}$$

In the general case we cannot assume the commutation relation of Eq. (3). It follows, that $\hat{c}_{\lambda_\mathrm{i}}$, $\hat{c}^\dagger_{\lambda_\mathrm{i}}$ are not, in general, ladder operators. We do, however, have the relationship of Eq. (4), and this is enough to prove

Lemma 2. *Given any retrodictively optimal measurement process with intial apparatus state $|\phi_\mathrm{ap}\rangle$ and retrodictive spatial resolution λ_i, let $\hat{c}_{\lambda_\mathrm{i}}$ be the operator defined by Eq. (6). Then*

$$\hat{c}_{\lambda_\mathrm{i}} |\psi \otimes \phi_\mathrm{ap}\rangle = 0$$

for every $|\psi\rangle \in \mathcal{H}_\mathrm{sy}$.

Proof. Given any normalised system state $|\psi\rangle$, let α, $\beta \in \mathbb{R}$ be the real and imaginary parts of $\langle \psi \otimes \phi_\mathrm{ap}| \, \hat{\epsilon}_\mathrm{Xi} \, \hat{\epsilon}_\mathrm{Pi} \, |\psi \otimes \phi_\mathrm{ap}\rangle$:

$$\langle \psi \otimes \phi_\mathrm{ap}| \, \hat{\epsilon}_\mathrm{Xi} \, \hat{\epsilon}_\mathrm{Pi} \, |\psi \otimes \phi_\mathrm{ap}\rangle = \alpha + i\beta \tag{7}$$

We have

$$\left(\alpha^2 + \beta^2\right)^{\frac{1}{2}} = \left|\langle \psi \otimes \phi_\mathrm{ap}| \, \hat{\epsilon}_\mathrm{Xi} \, \hat{\epsilon}_\mathrm{Pi} \, |\psi \otimes \phi_\mathrm{ap}\rangle\right| \le \left\|\hat{\epsilon}_\mathrm{Xi} \, |\psi \otimes \phi_\mathrm{ap}\rangle\right\| \, \left\|\hat{\epsilon}_\mathrm{Pi} \, |\psi \otimes \phi_\mathrm{ap}\rangle\right\| = \frac{\hbar}{2}$$

where

$$\left\|\hat{\epsilon}_\mathrm{Xi} \, |\psi \otimes \phi_\mathrm{ap}\rangle\right\| = \Big(\langle \psi \otimes \phi_\mathrm{ap}| \, \hat{\epsilon}^2_\mathrm{Xi} \, |\psi \otimes \phi_\mathrm{ap}\rangle\Big)^{\frac{1}{2}} = \frac{\lambda_\mathrm{i}}{\sqrt{2}}$$
$$\left\|\hat{\epsilon}_\mathrm{Pi} \, |\psi \otimes \phi_\mathrm{ap}\rangle\right\| = \Big(\langle \psi \otimes \phi_\mathrm{ap}| \, \hat{\epsilon}^2_\mathrm{Pi} \, |\psi \otimes \phi_\mathrm{ap}\rangle\Big)^{\frac{1}{2}} = \frac{\hbar}{\sqrt{2}\lambda_\mathrm{i}}$$

are the norms of the vectors $\hat{\epsilon}_\mathrm{Xi} \, |\psi \otimes \phi_\mathrm{ap}\rangle$, $\hat{\epsilon}_\mathrm{Pi} \, |\psi \otimes \phi_\mathrm{ap}\rangle$.

In view of Eq. (4) we also have

$$-i\hbar = \langle \psi \otimes \phi_\mathrm{ap}| \, [\hat{\epsilon}_\mathrm{Xi}, \hat{\epsilon}_\mathrm{Pi}] \, |\psi \otimes \phi_\mathrm{ap}\rangle = 2i\beta$$

Consequently, $\alpha = 0$ and $\beta = -\frac{\hbar}{2}$. We then have

$$\left|\langle \psi \otimes \phi_\mathrm{ap}| \, \hat{\epsilon}_\mathrm{Xi} \, \hat{\epsilon}_\mathrm{Pi} \, |\psi \otimes \phi_\mathrm{ap}\rangle\right| = \frac{\hbar}{2} = \left\|\hat{\epsilon}_\mathrm{Xi} \, |\psi \otimes \phi_\mathrm{ap}\rangle\right\| \, \left\|\hat{\epsilon}_\mathrm{Pi} \, |\psi \otimes \phi_\mathrm{ap}\rangle\right\|$$

Now it is generally true, in any Hilbert space, that two vectors $|\Psi_1\rangle$, $|\Psi_2\rangle$ having the property

$$\left|\langle \Psi_1 \mid \Psi_2\rangle\right| = \left\||\Psi_1\rangle\right\| \, \left\||\Psi_2\rangle\right\|$$

must be parallel. Hence

$$\hat{\epsilon}_\mathrm{Pi} \, |\psi \otimes \phi_\mathrm{ap}\rangle = \gamma \, \hat{\epsilon}_\mathrm{Xi} \, |\psi \otimes \phi_\mathrm{ap}\rangle$$

for some $\gamma \in \mathbb{C}$. Inserting this result into Eq. (7) we find

$$\gamma = -\frac{i\hbar}{\lambda_\mathrm{i}^2}$$

The claim follows. $\square$

Now let

$$\rho\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)=\int dx_{\mathrm{f}}\, dy_{\mathrm{f}1}\ldots dy_{\mathrm{f}n}\left|\langle x_{\mathrm{f}}, \mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}, y_{\mathrm{f}1}, \ldots, y_{\mathrm{f}n} \mid \psi \otimes \phi_{\mathrm{ap}}\rangle\right|^{2} \tag{8}$$

be the probability distribution for the final pointer positions. In this expression $|x_{\mathrm{f}}, \mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}, y_{\mathrm{f}1}, \ldots, y_{\mathrm{f}n}\rangle$ is the simultaneous eigenvector of the Heisenberg picture operators $\hat{x}_{\mathrm{f}}$, $\hat{\mu}_{\mathrm{Xf}}$, $\hat{\mu}_{\mathrm{Pf}}$, $\hat{y}_{\mathrm{f}j}$, with eigenvalues x_{f}, μ_{Xf}, μ_{Pf}, $y_{\mathrm{f}j}$. We continue to employ the notation and terminology of ref. [11]. Thus, $\hat{x}_{\mathrm{f}}$ is the final system position operator, $\hat{\mu}_{\mathrm{Xf}}$ and $\hat{\mu}_{\mathrm{Pf}}$ are the final pointer position operators, and the $\hat{y}_{\mathrm{f}j}$ represent the additional, internal degrees of freedom characterising the apparatus.

Let $\left|(x,p)_{\lambda_{\mathrm{i}}}\right\rangle \in \mathcal{H}_{\mathrm{sy}}$ be the state with wave function

$$\left\langle x' \,\middle|\, (x,p)_{\lambda_{\mathrm{i}}}\right\rangle=\left(\frac{1}{\pi\lambda_{\mathrm{i}}^{2}}\right)^{\frac{1}{4}} \exp\left[-\tfrac{1}{2\lambda_{\mathrm{i}}^{2}}(x'-x)^{2}+\tfrac{i}{\hbar}px'-\tfrac{i}{2\hbar}px\right] \tag{9}$$

and let

$$Q_{\lambda_{\mathrm{i}}}(x,p)=\frac{1}{h}\left|\left\langle (x,p)_{\lambda_{\mathrm{i}}} \middle| \psi\right\rangle\right|^{2} \tag{10}$$

be the initial system state Husimi function [13, 14]. We want to show

$$\rho\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)=Q_{\lambda_{\mathrm{i}}}\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)$$

for almost all μ_{Xf}, μ_{Pf} whenever the measurement is retrodictively optimal at spatial resolution λ_{i} ("almost all" being defined relative to ordinary Lebesgue measure on the plane). Our strategy will be to begin by showing that the two functions have the same moments:

$$\int d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}}\, \mu_{\mathrm{Xf}}^{n} \mu_{\mathrm{Pf}}^{m}\, \rho\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)=\int d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}}\, \mu_{\mathrm{Xf}}^{n} \mu_{\mathrm{Pf}}^{m}\, Q_{\lambda_{\mathrm{i}}}\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)$$

for every pair of non-negative integers n, m. Unfortunately we then face the difficulty, that although ρ and $Q_{\lambda_{\mathrm{i}}}$ are always defined, whatever the initial state of the system, the same is not true of their moments. This is because $\hat{x}_{\mathrm{i}}$, $\hat{p}_{\mathrm{i}}$, $\hat{\mu}_{\mathrm{Xf}}$, $\hat{\mu}_{\mathrm{Pf}}$ are unbounded operators. The way in which we will circumvent the difficulty is, first to prove the result on the assumption that $|\psi\rangle$ is in an appropriately chosen dense subspace of $\mathcal{H}_{\mathrm{sy}}$, and then to use a continuity argument to extend it to the case of arbitrary $|\psi\rangle$.

Let $\hat{a}_{\lambda_{\mathrm{i}}}$, $\hat{a}_{\lambda_{\mathrm{i}}}^{\dagger}$ be the ladder operators

$$\begin{aligned}
\hat{a}_{\lambda_{\mathrm{i}}} &=\frac{1}{\sqrt{2}}\left(\frac{1}{\lambda_{\mathrm{i}}} \hat{x}_{\mathrm{i}}+\frac{\lambda_{\mathrm{i}}}{\hbar} \hat{p}_{\mathrm{i}}\right) \\
\hat{a}_{\lambda_{\mathrm{i}}}^{\dagger} &=\frac{1}{\sqrt{2}}\left(\frac{1}{\lambda_{\mathrm{i}}} \hat{x}_{\mathrm{i}}-\frac{\lambda_{\mathrm{i}}}{\hbar} \hat{p}_{\mathrm{i}}\right)
\end{aligned} \tag{11}$$

and define number states $|n\rangle_{\lambda_{\mathrm{i}}} \in \mathcal{H}_{\mathrm{sy}}$ in the usual way, by the requirements

$$\hat{a}_{\lambda_{\mathrm{i}}}|0\rangle_{\lambda_{\mathrm{i}}}=0 \qquad {}_{\lambda_{\mathrm{i}}}\langle 0 \mid 0\rangle_{\lambda_{\mathrm{i}}}=1 \qquad |n\rangle_{\lambda_{\mathrm{i}}}=\frac{1}{\sqrt{n!}}\, \hat{a}_{\lambda_{\mathrm{i}}}^{\dagger\, n}\, |0\rangle_{\lambda_{\mathrm{i}}}$$

(with a slight abuse of notation we sometimes regard the operators $\hat{x}_{\mathrm{i}}$ and $\hat{p}_{\mathrm{i}}$ as acting on $\mathcal{H}_{\mathrm{sy}}$, and sometimes as acting on $\mathcal{H}_{\mathrm{sy}} \otimes \mathcal{H}_{\mathrm{ap}}$). We then define $\mathcal{F}_{\lambda_{\mathrm{i}}}$ to be the dense subspace of $\mathcal{H}_{\mathrm{sy}}$ consisting of all *finite* linear combinations of the vectors $|n\rangle_{\lambda_{\mathrm{i}}}$.

It is easily seen that $\mathcal{F}_{\lambda_{\rm i}}$ is in the domain of definition of every polynomial $f(\hat{x}_{\rm i}, \hat{p}_{\rm i})$. In particular, the integral

$$\int dxdp\, x^n p^m Q_{\lambda_{\rm i}}(x,p)$$

is defined and finite for all n, m whenever $Q_{\lambda_{\rm i}}$ is the Husimi function corresponding to a state in $\mathcal{F}_{\lambda_{\rm i}}$.

Now define the operators

$$\begin{aligned}\hat{b}_{\lambda_{\rm i}} &= \frac{1}{\sqrt{2}}\left(\frac{1}{\lambda_{\rm i}}\hat{\mu}_{\rm Xf} + \frac{i\lambda_{\rm i}}{\hbar}\hat{\mu}_{\rm Pf}\right)\\ \hat{b}^\dagger_{\lambda_{\rm i}} &= \frac{1}{\sqrt{2}}\left(\frac{1}{\lambda_{\rm i}}\hat{\mu}_{\rm Xf} - \frac{i\lambda_{\rm i}}{\hbar}\hat{\mu}_{\rm Pf}\right)\end{aligned}$$

These operators commute, and so they are certainly not ladder operators. We have

$$\hat{b}^\dagger_{\lambda_{\rm i}} = \hat{a}^\dagger_{\lambda_{\rm i}} + \hat{c}_{\lambda_{\rm i}} \tag{12}$$

where $\hat{c}_{\lambda_{\rm i}}$ and $\hat{a}^\dagger_{\lambda_{\rm i}}$ are the operators defined in Eqs. (6) and (11) respectively. Let $|\psi\rangle$ be any vector $\in \mathcal{F}_{\lambda_{\rm i}}$. Then $|\psi \otimes \phi_{\rm ap}\rangle$ is in the domain of $\hat{a}^\dagger_{\lambda_{\rm i}}$. It is also in the domain of $\hat{c}_{\lambda_{\rm i}}$ (the definition of a retrodictively optimal process tacitly assumes that $|\psi \otimes \phi_{\rm ap}\rangle$ is in the domain of $\hat{\epsilon}_{\rm Xi}$, $\hat{\epsilon}_{\rm Pi}$, and therefore in the domain of $\hat{c}_{\lambda_{\rm i}}$, for all $|\psi\rangle$). It is consequently in the domain of $\hat{b}^\dagger_{\lambda_{\rm i}}$. Moreover, in view of Lemma 2,

$$\hat{b}^\dagger_{\lambda_{\rm i}}|\psi \otimes \phi_{\rm ap}\rangle = \left(\hat{a}^\dagger_{\lambda_{\rm i}}|\psi\rangle\right) \otimes |\phi_{\rm ap}\rangle$$

where $\hat{a}^\dagger_{\lambda_{\rm i}}|\psi\rangle$ also $\in \mathcal{F}_{\lambda_{\rm i}}$. Iterating the argument we conclude that $|\psi \otimes \phi_{\rm ap}\rangle$ is in the domain of $\hat{b}^{\dagger\, n}_{\lambda_{\rm i}}$ and

$$\hat{b}^{\dagger\, n}_{\lambda_{\rm i}}|\psi \otimes \phi_{\rm ap}\rangle = \left(\hat{a}^{\dagger\, n}_{\lambda_{\rm i}}|\psi\rangle\right) \otimes |\phi_{\rm ap}\rangle$$

for every non-negative integer n. Taking adjoints gives

$$\langle\psi \otimes \phi_{\rm ap}|\,\hat{b}^{m}_{\lambda_{\rm i}} = \left(\langle\psi|\,\hat{a}^{m}_{\lambda_{\rm i}}\right) \otimes \langle\phi_{\rm ap}|$$

for all m. Consequently,

$$\big\langle\psi \otimes \phi_{\rm ap}\big|\,\hat{b}^{m}_{\lambda_{\rm i}}\hat{b}^{\dagger\, n}_{\lambda_{\rm i}}\big|\psi \otimes \phi_{\rm ap}\big\rangle = \big\langle\psi\big|\,\hat{a}^{m}_{\lambda_{\rm i}}\hat{a}^{\dagger\, n}_{\lambda_{\rm i}}\big|\psi\big\rangle$$

Now

$$\big\langle\psi \otimes \phi_{\rm ap}\big|\,\hat{b}^{m}_{\lambda_{\rm i}}\hat{b}^{\dagger\, n}_{\lambda_{\rm i}}\big|\psi \otimes \phi_{\rm ap}\big\rangle = \int d\mu_{\rm Xf}d\mu_{\rm Pf}\, z^{m}_{\lambda_{\rm i}} z^{*\, n}_{\lambda_{\rm i}}\, \rho\left(\mu_{\rm Xf}, \mu_{\rm Pf}\right)$$

where ρ is the distribution of final pointer positions, as defined in Eq. (8), and $z_{\lambda_{\rm i}}$ is the complex coordinate

$$z_{\lambda_{\rm i}} = \frac{1}{\sqrt{2}}\left(\frac{1}{\lambda_{\rm i}}\mu_{\rm Xf} + \frac{i\lambda_{\rm i}}{\hbar}\mu_{\rm Pf}\right) \tag{13}$$

Also [14]

$$\big\langle\psi\big|\,\hat{a}^{m}_{\lambda_{\rm i}}\hat{a}^{\dagger\, n}_{\lambda_{\rm i}}\big|\psi\big\rangle = \int d\mu_{\rm Xf}d\mu_{\rm Pf}\, z^{m}_{\lambda_{\rm i}} z^{*\, n}_{\lambda_{\rm i}}\, Q_{\lambda_{\rm i}}\left(\mu_{\rm Xf}, \mu_{\rm Pf}\right) \tag{14}$$

where $Q_{\lambda_{\mathrm{i}}}$ is the initial system state Husimi function, as defined in Eq. (10). Therefore

$$\int d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}}\, z_{\lambda_{\mathrm{i}}}^m z_{\lambda_{\mathrm{i}}}^{*\,n}\, \rho\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) = \int d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}}\, z_{\lambda_{\mathrm{i}}}^m z_{\lambda_{\mathrm{i}}}^{*\,n}\, Q_{\lambda_{\mathrm{i}}}\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)$$

for all n, m. It follows that

$$\int d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}}\, f(z_{\lambda_{\mathrm{i}}}, z_{\lambda_{\mathrm{i}}}^*)\, \rho\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) = \int d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}}\, f(z_{\lambda_{\mathrm{i}}}, z_{\lambda_{\mathrm{i}}}^*)\, Q_{\lambda_{\mathrm{i}}}\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)$$

for every polynomial f. In particular

$$\int d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}}\, \mu_{\mathrm{Xf}}^m \mu_{\mathrm{Pf}}^n\, \rho\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) = \int d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}}\, \mu_{\mathrm{Xf}}^m \mu_{\mathrm{Pf}}^n\, Q_{\lambda_{\mathrm{i}}}\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) \tag{15}$$

for all m, n.

At this stage one needs to be careful. It is tempting to suppose, that two probability measures which have the same moments must be equal. In fact, this inference is not always justified (see Reed and Simon [19], vol. 2). However, it is justified here, as we show in the Appendix. Consequently

$$\rho\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) = Q_{\lambda_{\mathrm{i}}}\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) \tag{16}$$

for almost all μ_{Xf}, μ_{Pf} whenever the initial system state $|\psi\rangle$ is in the space $\mathcal{F}_{\lambda_{\mathrm{i}}}$.

It remains for us to show that the distributions are equal in the case of arbitrary $|\psi\rangle \in \mathcal{H}_{\mathrm{sy}}$. We will do this by using a continuity argument.

Choose a sequence $|\psi_n\rangle \in \mathcal{F}_{\lambda_{\mathrm{i}}}$ converging to $|\psi\rangle$. Let $Q_{\lambda_{\mathrm{i}},n}$ be the Husimi function, and ρ_n the distribution of measured values corresponding to $|\psi_n\rangle$. Let $Q_{\lambda_{\mathrm{i}}}$ be the Husimi function, and ρ the distribution of measured values corresponding to $|\psi\rangle$.

We have, as an immediate consequence of the definition, Eq. (10),

$$Q_{\lambda_{\mathrm{i}}}\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) = \lim_{n\to\infty}\left(Q_{\lambda_{\mathrm{i}},n}\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)\right) \tag{17}$$

for all μ_{Xf}, μ_{Pf}.

On the other hand, it is not generally true that ρ_n converges pointwise to ρ. It does, however, contain a subsequence which converges pointwise almost everywhere. In fact, let $\mathcal{L}_1$ be the Banach space consisting of all integrable functions on $\mathbb{R}^2$, with norm

$$\|f\|_1 = \int d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}} \left|f\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)\right|$$

We have

$$\begin{aligned}
\|\rho - \rho_n\|_1 = \int d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}} & \left| \int dx_{\mathrm{f}} dy_{\mathrm{f1}} \dots y_{\mathrm{f}n} \, \Big(\left|\langle x_{\mathrm{f}}, \mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}, y_{\mathrm{f1}}, \dots, y_{\mathrm{f}n} \mid \psi \otimes \phi_{\mathrm{ap}}\rangle\right|^2 \right. \\
& \left. \qquad - \left|\langle x_{\mathrm{f}}, \mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}, y_{\mathrm{f1}}, \dots, y_{\mathrm{f}n} \mid \psi_n \otimes \phi_{\mathrm{ap}}\rangle\right|^2 \Big) \right| \\
\le & \left\| |\psi \otimes \phi_{\mathrm{ap}}\rangle - |\psi_n \otimes \phi_{\mathrm{ap}}\rangle \right\| \Big(\left\| |\psi \otimes \phi_{\mathrm{ap}}\rangle \right\| + \left\| |\psi \otimes \phi_{\mathrm{ap}}\rangle \right\| \Big) \\
\to & \ 0
\end{aligned}$$

We see from this that $\rho_n \to \rho$ in the topology of $\mathcal{L}_1$. We may therefore use the Riesz-Fisher theorem (Reed and Simon [19], vol. 1) to deduce that it contains a subsequence ρ_{n_r} such that

$$\rho\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) = \lim_{r\to\infty}\left(\rho_{n_r}\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)\right)$$

for almost all μ_{Xf}, μ_{Pf}. In view of this result, Eq. (17), and the fact that

$$\rho_{n_r}(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}) = \Omega_{\lambda_{\mathrm{i}}, n_r}(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}})$$

for all r and almost all μ_{Xf}, μ_{Pf} we deduce that

$$\rho(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}) = \Omega_{\lambda_{\mathrm{i}}}(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}})$$

for almost all μ_{Xf}, μ_{Pf}.

3. Predictively Optimal Measurements

We will say that a simultaneous measurement process of the kind defined in ref. [11] is *predictively optimal* if the product of predictive errors is minimised:

$$\Delta_{\mathrm{ef}} x \, \Delta_{\mathrm{ef}} p = \frac{\hbar}{2} \tag{18}$$

In view of the commutation relation

$$[\hat{\epsilon}_{\mathrm{Xf}}, \hat{\epsilon}_{\mathrm{Pf}}] = i\hbar \tag{19}$$

there is no need to impose the condition, that the measurement be predictively unbiased as a separate requirement: it is a consequence of the condition of Eq. (18).

Eqs. (18) and (19) together imply

$$\left\langle \psi \otimes \phi_{\mathrm{ap}} \right| \hat{\epsilon}_{\mathrm{Xf}}^2 \left| \psi \otimes \phi_{\mathrm{ap}} \right\rangle \left\langle \psi \otimes \phi_{\mathrm{ap}} \right| \hat{\epsilon}_{\mathrm{Pf}}^2 \left| \psi \otimes \phi_{\mathrm{ap}} \right\rangle = \frac{\hbar^2}{4}$$

for every normalised $|\psi\rangle \in \mathcal{H}_{\mathrm{sy}}$. By an argument which parallels the proof of Lemma 1 we infer that there exists a fixed number λ_{f} such that

$$\left\langle \psi \otimes \phi_{\mathrm{ap}} \right| \hat{\epsilon}_{\mathrm{Xf}}^2 \left| \psi \otimes \phi_{\mathrm{ap}} \right\rangle = \frac{\lambda_{\mathrm{f}}^2}{2}$$
$$\left\langle \psi \otimes \phi_{\mathrm{ap}} \right| \hat{\epsilon}_{\mathrm{Pf}}^2 \left| \psi \otimes \phi_{\mathrm{ap}} \right\rangle = \frac{\hbar^2}{2\lambda_{\mathrm{f}}^2}$$

for every normalised $|\psi\rangle \in \mathcal{H}_{\mathrm{sy}}$. It is then straightforward to show that

$$\hat{d}_{\lambda_{\mathrm{f}}} \left| \psi \otimes \phi_{\mathrm{ap}} \right\rangle = 0 \tag{20}$$

for all $|\psi\rangle \in \mathcal{H}_{\mathrm{sy}}$, where $\hat{d}_{\lambda_{\mathrm{f}}}$ is the annihilation operator

$$\hat{d}_{\lambda_{\mathrm{f}}} = \frac{1}{\sqrt{2}} \left(\frac{1}{\lambda_{\mathrm{f}}} \hat{\epsilon}_{\mathrm{Xf}} + \frac{i\lambda_{\mathrm{f}}}{\hbar} \hat{\epsilon}_{\mathrm{Pf}} \right)$$

Since $\hat{\epsilon}_{\mathrm{Xf}}$, $\hat{\epsilon}_{\mathrm{Pf}}$ are canonically conjugate there exist kets $|\epsilon_{\mathrm{Xf}}, \mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}, y_{\mathrm{f}1}, \dots, y_{\mathrm{f}n}\rangle_{\epsilon}$ which are simultaneous eigenvectors of the operators $\hat{\epsilon}_{\mathrm{Xf}}$, $\hat{\mu}_{\mathrm{Xf}}$, $\hat{\mu}_{\mathrm{Pf}}$, $\hat{y}_{\mathrm{f}j}$, and which have the property

$${}_{\epsilon}\langle \epsilon_{\mathrm{Xf}}, \mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}, y_{\mathrm{f}1}, \dots, y_{\mathrm{f}n} | \, \hat{\epsilon}_{\mathrm{Pf}} \, |\Psi\rangle = -i\hbar \frac{\partial}{\partial \epsilon_{\mathrm{Xf}}} \, {}_{\epsilon}\langle \epsilon_{\mathrm{Xf}}, \mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}, y_{\mathrm{f}1}, \dots, y_{\mathrm{f}n} \mid \Psi \rangle \tag{21}$$

for all $|\Psi\rangle \in \mathcal{H}_{\mathrm{sy}} \otimes \mathcal{H}_{\mathrm{ap}}$. In view of Eq. (20) we then have

$$\left(\frac{1}{\lambda_{\mathrm{f}}} \epsilon_{\mathrm{Xf}} + \lambda_{\mathrm{f}} \frac{\partial}{\partial \epsilon_{\mathrm{Xf}}} \right) {}_{\epsilon}\langle \epsilon_{\mathrm{Xf}}, \mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}, y_{\mathrm{f}1}, \dots, y_{\mathrm{f}n} \mid \psi \otimes \phi_{\mathrm{ap}} \rangle = 0$$

for all $|\psi\rangle \in \mathcal{H}_{\text{sy}}$. Solving this equation we find

$$\begin{aligned}&{}_{\epsilon}\langle \epsilon_{\text{Xf}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n} \,|\, \psi \otimes \phi_{\text{ap}} \rangle \\ &\quad = \left(\frac{1}{\pi \lambda_{\text{f}}^2}\right)^{\frac{1}{4}} \exp\left[-\tfrac{1}{2\lambda_{\text{f}}^2} \epsilon_{\text{Xf}}^2\right] \Phi\left(\mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n}\right)\end{aligned} \tag{22}$$

where Φ is an arbitrary normalised function.

There also exist kets $|x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n}\rangle_x$ which are simultaneous eigenvectors of the operators $\hat{x}_{\text{f}}$, $\hat{\mu}_{\text{Xf}}$, $\hat{\mu}_{\text{Pf}}$, $\hat{y}_{\text{f}j}$ with the property

$${}_{x}\langle x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n} |\, \hat{p}_{\text{f}}\, |\Psi\rangle = -i\hbar \frac{\partial}{\partial x_{\text{f}}}\, {}_{x}\langle x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n} \,|\, \Psi\rangle \tag{23}$$

for all $|\Psi\rangle \in \mathcal{H}_{\text{sy}} \otimes \mathcal{H}_{\text{ap}}$. In view of the defining relation $\hat{\epsilon}_{\text{Xf}} = \hat{\mu}_{\text{Xf}} - \hat{x}_{\text{f}}$ we must have

$$\begin{aligned}&|x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n}\rangle_x \\ &\quad = e^{-i\chi(x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n})} \,|\mu_{\text{Xf}} - x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n}\rangle_{\epsilon}\end{aligned} \tag{24}$$

where $e^{-i\chi(x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n})}$ is a phase. In view of Eqs. (21) and (23) we must then have

$$\begin{aligned}&{}_{x}\langle x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n} |\, \hat{p}_{\text{f}}\, |\Psi\rangle \\ &\quad = -i\hbar \frac{\partial}{\partial x_{\text{f}}} \left(e^{i\chi(x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n})}\, {}_{\epsilon}\langle \mu_{\text{Xf}} - x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n} \,|\, \Psi\rangle \right)\end{aligned}$$

and

$$\begin{aligned}&{}_{x}\langle x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n} |\, \hat{p}_{\text{f}}\, |\Psi\rangle \\ &\quad = e^{i\chi(x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n})}\, {}_{\epsilon}\langle \mu_{\text{Xf}} - x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n} |\, \hat{\mu}_{\text{Pf}} - \hat{\epsilon}_{\text{Pf}}\, |\Psi\rangle \\ &\quad = e^{i\chi(x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n})} \\ &\qquad\quad \times \left(\mu_{\text{Pf}} + i\hbar \frac{\partial}{\partial \hat{\epsilon}_{\text{Xf}}}\right) {}_{\epsilon}\langle \hat{\epsilon}_{\text{Xf}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n} \,|\, \Psi\rangle\big|_{\hat{\epsilon}_{\text{Xf}} = \mu_{\text{Xf}} - x_{\text{f}}}\end{aligned}$$

for all $|\Psi\rangle \in \mathcal{H}_{\text{sy}} \otimes \mathcal{H}_{\text{ap}}$. Hence

$$\hbar \frac{\partial}{\partial x_{\text{f}}} \chi\left(x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n}\right) = \mu_{\text{Pf}}$$

which implies

$$\chi\left(x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n}\right) = \frac{1}{\hbar} \mu_{\text{Pf}} x_{\text{f}} + \chi_0\left(\mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n}\right)$$

where χ_0 is an arbitrary function. Using this result and Eq. (24) in Eq. (22) we deduce, that the final state wave function can be written

$$\begin{aligned}&{}_{x}\langle x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n} \,|\, \psi \otimes \phi_{\text{ap}} \rangle \\ &\quad = \left(\frac{1}{\pi \lambda_{\text{f}}^2}\right)^{\frac{1}{4}} \exp\left[-\tfrac{1}{2\lambda_{\text{f}}^2} \left(\mu_{\text{Xf}} - x_{\text{f}}\right)^2 + \tfrac{i}{\hbar} \mu_{\text{Pf}} x_{\text{f}} + i\chi_0\left(\mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n}\right)\right] \\ &\qquad\qquad \times \Phi\left(\mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n}\right)\end{aligned}$$

In terms of the state $\left|(\mu_{\text{Xf}}, \mu_{\text{Pf}})_{\lambda_{\text{f}}}\right\rangle$ defined in Eq. (9) this becomes

$${}_{x}\langle x_{\text{f}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n} \,|\, \psi \otimes \phi_{\text{ap}} \rangle = \left\langle x_{\text{f}} \,\middle|\, (\mu_{\text{Xf}}, \mu_{\text{Pf}})_{\lambda_{\text{f}}} \right\rangle \Phi'\left(\mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{\text{f}1}, \dots, y_{\text{f}n}\right)$$

where

$$\begin{aligned}&\Phi'\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}, y_{\mathrm{f}1}, \dots, y_{\mathrm{f}n}\right)\\ &\qquad = \exp\bigl[i\chi_0\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}, y_{\mathrm{f}1}, \dots, y_{\mathrm{f}n}\right) + \tfrac{i}{2\hbar}\mu_{\mathrm{Pf}}\mu_{\mathrm{Xf}}\bigr]\,\Phi\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}, y_{\mathrm{f}1}, \dots, y_{\mathrm{f}n}\right)\end{aligned}$$

The distribution of measured values $\rho\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)$ can be written in terms of Φ':

$$\rho\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) = \int dy_{\mathrm{f}1} \dots y_{\mathrm{f}n}\, \left|\Phi'\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}, y_{\mathrm{f}1}, \dots, y_{\mathrm{f}n}\right)\right|^2$$

Suppose, now, that the pointer positions are found to be in the region $\mathcal{R} \subseteq \mathbb{R}^2$. Let $\hat{\rho}_{\mathrm{sy}}$ be the reduced density matrix describing the state of the system immediately afterwards. Then

$$\begin{aligned}\langle x_{\mathrm{f}1}|\, \hat{\rho}_{\mathrm{sy}}\, |x_{\mathrm{f}2}\rangle = \frac{1}{p_{\mathcal{R}}}\int_{\mathcal{R}\times\mathbb{R}^n} & d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}} dy_{\mathrm{f}1} \dots dy_{\mathrm{f}n}\, \left|\Phi'\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}, y_{\mathrm{f}1}, \dots, y_{\mathrm{f}n}\right)\right|^2 \\ & \times \bigl\langle x_{\mathrm{f}1}\,\big|\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)_{\lambda_{\mathrm{f}}}\bigr\rangle\bigl\langle\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)_{\lambda_{\mathrm{f}}}\big|\, x_{\mathrm{f}2}\bigr\rangle\end{aligned}$$

where $p_{\mathcal{R}}$ is the probability of finding $\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) \in \mathcal{R}$:

$$p_{\mathcal{R}} = \int_{\mathcal{R}} d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}}\, \rho\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)$$

Hence

$$\hat{\rho}_{\mathrm{sy}} = \frac{1}{p_{\mathcal{R}}}\int_{\mathcal{R}} d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}}\, \rho\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) \bigl|\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)_{\lambda_{\mathrm{f}}}\bigr\rangle\bigl\langle\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)_{\lambda_{\mathrm{f}}}\bigr|$$

On the other hand

$$\hat{\rho}_{\mathrm{sy}} = \int d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}}\, P_{\lambda_{\mathrm{f}}}\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) \bigl|\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)_{\lambda_{\mathrm{f}}}\bigr\rangle\bigl\langle\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)_{\lambda_{\mathrm{f}}}\bigr|$$

where $P_{\lambda_{\mathrm{f}}}$ is the anti-Husimi function (or P-function) [14, 18] describing the final state of the system. Comparing these expressions we see

$$P_{\lambda_{\mathrm{f}}}\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) = \begin{cases}\frac{1}{p_{\mathcal{R}}}\rho\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) & \text{if } \left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) \in \mathcal{R} \\ 0 & \text{otherwise}\end{cases} \tag{25}$$

If $\mathcal{R}$ is a sufficiently small region centred on the point $\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)$ the system is approximately in the state $\bigl|\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)_{\lambda_{\mathrm{f}}}\bigr\rangle$ after the measurement:

$$\hat{\rho}_{\mathrm{sy}} \approx \bigl|\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)_{\lambda_{\mathrm{f}}}\bigr\rangle\bigl\langle\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right)_{\lambda_{\mathrm{f}}}\bigr|$$

Eq. (25) shows that the effect of a predictively optimal measurement process is to leave the system in a state for which $P_{\lambda_{\mathrm{f}}}$ is a probability density function. Such states are, of course, exceptional. In many cases, $P_{\lambda_{\mathrm{f}}}$ is not even defined as a tempered distribution [14].

4. The Interpretation of the Husimi Function

The result proved in Section 2 shows that there is a certain analogy between the Husimi function and the x-space probability density function $\left|\langle x \mid \psi\rangle\right|^2$. To see this let us examine just what is meant by the statement, that $\left|\langle x \mid \psi\rangle\right|^2 \delta x$ represents the probability of finding the position to lie in the interval $(x, x + \delta x)$.

Consider a measurement of x only. For the sake of simplicity suppose that the measuring apparatus has only one degree of freedom, corresponding to the single pointer observable $\hat{\mu}_{\mathrm{X}}$ (the argument which follows does not depend on this assumption, however). Let $|\psi\rangle$ and $|\phi_{\mathrm{ap}}\rangle$ be the initial states of the system and

apparatus respectively, and let $\hat{U}$ be the unitary evolution operator describing the measurement interaction. Let $\hat{x}_{\mathrm{i}} = \hat{x}$ and $\hat{\mu}_{\mathrm{Xf}} = \hat{U}^{\dagger} \mu_{\mathrm{Xf}} \hat{U}$ be the Heisenberg picture operators describing the initial position of the system and final position of the pointer respectively. Let $\hat{\epsilon}_{\mathrm{Xi}} = \hat{\mu}_{\mathrm{Xf}} - \hat{x}_{\mathrm{i}}$ be the retrodictive error operator.

The final state wave function can be written (in the Schrödinger picture)

$$\left\langle x, \mu_{\mathrm{X}} \right| \hat{U} \left| \psi \otimes \phi_{\mathrm{ap}} \right\rangle = \int dx' \, K\left(x, \mu_{\mathrm{X}}; x'\right) \left\langle x' \mid \psi \right\rangle$$

for some kernel K. The probability distribution describing the result of the measurement then takes the form

$$\rho\left(\mu_{\mathrm{X}}\right) = \int dx \left| \int dx' \, K\left(x, \mu_{\mathrm{X}}; x'\right) \left\langle x' \mid \psi \right\rangle \right|^{2} \tag{26}$$

After a certain amount of algebra one also finds

$$\left\langle \psi \otimes \phi_{\mathrm{ap}} \right| \hat{\epsilon}_{\mathrm{Xi}}^{2} \left| \psi \otimes \phi_{\mathrm{ap}} \right\rangle = \int dx d\mu_{\mathrm{X}} \left| \int dx' \left(\mu_{\mathrm{X}} - x'\right) K\left(x, \mu_{\mathrm{X}}; x'\right) \left\langle x' \mid \psi \right\rangle \right|^{2} \tag{27}$$

Suppose that $\Delta_{\mathrm{ei}} x = 0$. Then we see from Eq. (27) that K must take the form

$$K\left(x, \mu_{\mathrm{X}}; x'\right) = f\left(x, \mu_{\mathrm{X}}\right) \delta\left(\mu_{\mathrm{X}} - x'\right)$$

for some function f. The unitarity of $\hat{U}$ means that f must satisfy

$$\int dx \, \left| f\left(x, \mu_{\mathrm{X}}\right) \right|^{2} = 1$$

Using these results in Eq. (26) we find

$$\rho\left(\mu_{\mathrm{X}}\right) = \left| \left\langle \mu_{\mathrm{X}} \mid \psi \right\rangle \right|^{2}$$

whenever the measurement is perfectly accurate for the purposes of retrodiction.

Suppose, on the other hand, that $\Delta_{\mathrm{ei}} x > 0$. Then $\rho\left(\mu_{\mathrm{X}}\right)$ will not generally coincide with the function $\left| \left\langle \mu_{\mathrm{X}} \mid \psi \right\rangle \right|^{2}$. If $\Delta_{\mathrm{ei}} x$ is small compared with the de Broglie wavelength, then we see from Eqs. (26) and (27) that $\rho\left(\mu_{\mathrm{X}}\right) \approx \left| \left\langle \mu_{\mathrm{X}} \mid \psi \right\rangle \right|^{2}$. Otherwise, we do not expect the two functions even to be approximately equal.

Although one may possibly approach, one does not expect actually to achieve the limit of perfect accuracy. It follows, that one does not expect the function $\left| \left\langle \mu_{\mathrm{X}} \mid \psi \right\rangle \right|^{2}$ to describe the outcome of any practically realisable measurement of position.

This being so what, exactly, is the significance of the function $\left| \left\langle \mu_{\mathrm{X}} \mid \psi \right\rangle \right|^{2}$? In the first place, it serves as a standard of comparison, against which the outcome of experimentally realisable measurements can be judged: in the sense, that the better the measurement, the more closely does the function $\left| \left\langle \mu_{\mathrm{X}} \mid \psi \right\rangle \right|^{2}$ approximate the distribution of actual results.

In the second place, we see from Eq. (26) that the outcome of a real measurement of position depends, not only on the state of the system, *via* the function $\left\langle x' \mid \psi \right\rangle$, but also on the details of the measurement process, *via* the function $K\left(x, \mu_{\mathrm{X}}; x'\right)$. In the limit of perfect retrodictive accuracy, however, the dependence on the apparatus (as represented by the kernel K) disappears, and the distribution of results is determined solely by the state of the system (as represented by the vector $\left| \psi \right\rangle$). $\left| \left\langle \mu_{\mathrm{X}} \mid \psi \right\rangle \right|^{2}$ does, so to speak, represent the *intrinsic* distribution of position, independent of any properties specific to the particular measuring instrument employed.

In a real measurement, by contrast, the outcome is (in a manner of speaking) contaminated by instrumental contributions, which one may try to reduce, but can never entirely eliminate.

One typically regards the function $|\langle \mu_{\mathrm{X}} \mid \psi \rangle|^2$ simply, and without qualification, as *the* x-space probability distribution. It owes this canonical status to the two features just mentioned. The result proved in Section 2 shows that the Husimi function has analogous features. It describes the outcome of those measurements which are retrodictively optimal, or “best”. It is otherwise independent of the details of the particular process considered. It might therefore be regarded as the canonical probability distribution for position and momentum.

In classical mechanics one has the concept of the “actual” distribution describing an ensemble of identically prepared systems. Quantum mechanics contains no precise analogue for this concept (unless one adopts a “hidden-variables” interpretation [20]). Nevertheless, the result proved in Section 2 shows that there are certain resemblances between the Husimi function and the classical distribution. The Husimi function is clearly not the same as the classical distribution. However, one might reasonably argue that it is the closest that quantum mechanics allows us to get to the concept of a “real” or “objective” phase space probability distribution.

Appendix. Proof of Equation (16)

Rather than working in terms of the functions ρ, $Q_{\lambda_{\mathrm{i}}}$ it will be convenient, instead, to work in terms of the measures

$$\begin{aligned} d\mu_\rho &= \rho\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}} \\ d\mu_Q &= Q_{\lambda_{\mathrm{i}}}\left(\mu_{\mathrm{Xf}}, \mu_{\mathrm{Pf}}\right) d\mu_{\mathrm{Xf}} d\mu_{\mathrm{Pf}} \end{aligned}$$

We have from Eqs. (14) and (15)

$$\int d\mu_\rho \, |z_{\lambda_{\mathrm{i}}}|^{2n} = \int d\mu_Q \, |z_{\lambda_{\mathrm{i}}}|^{2n} = \langle \psi | \, \hat{a}^n_{\lambda_{\mathrm{i}}} \hat{a}^{\dagger\, n}_{\lambda_{\mathrm{i}}} \, |\psi\rangle \tag{28}$$

where $z_{\lambda_{\mathrm{i}}}$ is the complex co-ordinate defined in Eq. (13). Our strategy will be, first to establish a bound on the rate at which these quantities grow with increasing n, and then to use this to show that the measures μ_ρ, μ_Q have the same Fourier transform.

$|\psi\rangle$ is in the subspace $\mathcal{F}_{\lambda_{\mathrm{i}}}$. It can therefore be written

$$|\psi\rangle = \sum_{r=0}^{l} c_r \, |r\rangle_{\lambda_{\mathrm{i}}}$$

for some integer l. Hence

$$\langle \psi | \, \hat{a}^n_{\lambda_{\mathrm{i}}} \hat{a}^{\dagger\, n}_{\lambda_{\mathrm{i}}} \, |\psi\rangle = \sum_{r=0}^{l} \frac{(n+r)!}{r!} \, |c_r|^2 \leq \frac{(n+l)!}{l!}$$

Let μ stand for either of the measures μ_ρ, μ_Q. In view of the inequality just proved, Eq. (28) and the fact

$$|z_{\lambda_{\mathrm{i}}}|^{2n+1} \leq \frac{1}{2}\left(|z_{\lambda_{\mathrm{i}}}|^{2n} + |z_{\lambda_{\mathrm{i}}}|^{2n+2}\right)$$

we have

$$\int d\mu \, |z_{\lambda_{\mathrm{i}}}|^n \leq \frac{\Gamma\left(\frac{1}{2}n + l + \frac{3}{2}\right)}{\Gamma(l+1)}$$

for every non-negative integer n. Hence

$$\sum_{n=0}^{\infty} \frac{1}{n!} \int d\mu \, \left| \beta z_{\lambda_{\rm i}} + \gamma z_{\lambda_{\rm i}}^* \right|^n < \infty$$

for all β, $\gamma \in \mathbb{C}$. It follows that the functions $e^{|\beta z_{\lambda_{\rm i}} + \gamma z_{\lambda_{\rm i}}^*|}$ and $e^{\beta z_{\lambda_{\rm i}} + \gamma z_{\lambda_{\rm i}}^*}$ are μ-integrable. We may therefore use Lebesgue's dominated convergence theorem (Reed and Simon [19], vol. 1) to infer

$$\begin{aligned} \int d\mu_\rho \, \exp\left[\beta z_{\lambda_{\rm i}} + \gamma z_{\lambda_{\rm i}}^*\right] &= \lim_{N \to \infty} \left(\sum_{n=0}^{N} \frac{1}{n!} \int d\mu_\rho \, \left(\beta z_{\lambda_{\rm i}} + \gamma z_{\lambda_{\rm i}}^*\right)^n \right) \\ &= \lim_{N \to \infty} \left(\sum_{n=0}^{N} \frac{1}{n!} \int d\mu_Q \, \left(\beta z_{\lambda_{\rm i}} + \gamma z_{\lambda_{\rm i}}^*\right)^n \right) \\ &= \int d\mu_Q \, \exp\left[\beta z_{\lambda_{\rm i}} + \gamma z_{\lambda_{\rm i}}^*\right] \end{aligned}$$

for all β, $\gamma \in \mathbb{C}$. Consequently

$$\int d\mu_\rho \, \exp\left[i\left(k_{\rm X}\mu_{\rm Xf} + k_{\rm P}\mu_{\rm Pf}\right)\right] = \int d\mu_Q \, \exp\left[i\left(k_{\rm X}\mu_{\rm Xf} + k_{\rm P}\mu_{\rm Pf}\right)\right]$$

for all $k_{\rm X}$, $k_{\rm P} \in \mathbb{R}$. Inverting the Fourier transforms we deduce

$$\mu_\rho = \mu_Q$$

References

[1] E. Arthurs and J. L. Kelly, Bell Syst. Tech. J. **44**, 725 (1965); S. L. Braunstein, C. M. Caves and G. J. Milburn, Phys. Rev. A **43**, 1153 (1991); S. Stenholm, Ann. Phys. (N.Y.) **218**, 233 (1992).

[2] J. J. Halliwell, Phys. Rev. D **46**, 1610 (1992).

[3] U. Leonhardt and H. Paul, J. Mod. Opt. **40**, 1745 (1993); U. Leonhardt and H. Paul, Prog. Quant. Electr. **19**, 89 (1995).

[4] P. Törma, S. Stenholm and I. Jex, Phys. Rev. A **52**, 4812 (1995).

[5] W. L. Power, S. M. Tan and M. Wilkens, J. Mod. Opt., **44**, 2591 (1997).

[6] K. Wódkiewicz, Phys. Rev. Lett. **52**, 1064 (1984). Also see K. Wódkiewicz, Phys. Lett. A **115**, 304 (1986); *ibid.* **124**, 207 (1987); D. Burak and K. Wódkiewicz, Phys. Rev. A **46**, 2744 (1992); B-G. Englert and K. Wódkiewicz, Phys. Rev. A **51**, R2661 (1995); K. Banaszek and K. Wódkiewicz, Phys. Rev. A **55**, 3117 (1997).

[7] M. Ban, Int. J. Theor. Phys. **36**, 2583 (1997).

[8] U. Leonhardt, *Measuring the Quantum State of Light* (Cambridge University Press, Cambridge, 1997).

[9] P. Busch, M. Grabowski and P. J. Lahti, *Operational Quantum Physics* (Springer-Verlag, Berlin, 1995).

[10] D. M. Appleby, report no. QMW-PH-98-11, xxx.lanl.gov quant-ph/9803046. To be published in Int. J. Theor. Phys.

[11] D. M. Appleby, report no. QMW-PH-98-13, xxx.lanl.gov quant-ph/9803051. Submitted to Phys. Rev. A.

[12] V. B. Braginsky and F. Ya. Khalili, *Quantum Measurement,* edited by K. S. Thorne (Cambridge University Press, Cambridge, 1992).

[13] K. Husimi, Proc. Phys. Math. Soc. Jpn. **22**, 264 (1940).

[14] M. Hillery, R. F. O'Connell, M. O. Scully and E. P. Wigner, Phys. Rep. **106**, 121 (1984); H. W. Lee, Phys. Rep. **259**, 147 (1995).

[15] S. T. Ali and E. Prugovečki, J. Math. Phys. **18**, 219 (1977). Also see E. Prugovečki, J. Math. Phys. **17**, 517 (1976); *ibid.* **17**, 1673 (1976); Ann. Phys. (N.Y.) **110**, 102 (1978); A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982);

and E. Prugovečki, *Stochastic Quantum Mechanics and Quantum Space-time* (Reidel, Dordrecht, 1984).

[16] D. M. Appleby, report no. QMW-PH-98-12, xxx.lanl.gov quant-ph/9803047. Submitted to J. Phys. A.

[17] J. Uffink, Int. J. Theor. Phys. **33**, 199 (1994).

[18] R. J. Glauber, Phys. Rev. **131**, 2766 (1963); E. C. G. Sudarshan, Phys. Rev. Lett. **10**, 277 (1963).

[19] M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, vols 1–4 (Academic Press, New York, 1975–80).

[20] D. Bohm and B. J. Hiley, *The Undivided Universe* (Routledge, London, 1993); P. R. Holland, *The Quantum Theory of Motion* (Cambridge University Press, Cambridge, 1993); and references cited therein.